Journal of Multivariate Analysis 101 (2010) 583–593

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Near-exact distributions for the independence and sphericity likelihood ratio test statistics

Carlos A. Coelho *, Filipe J. Marques

Mathematics Department, Faculty of Sciences and Technology, The New University of Lisbon, Portugal

ARTICLE INFO

Article history:
Received 11 December 2007
Available online 29 September 2009

AMS subject classifications:
62E20
62H05
62H10
62H15

Keywords:
Wilks lambda statistic
Independence test
Sphericity test
Generalized Near-Integer Gamma distribution
Mixtures

ABSTRACT

In this paper we show how, based on a decomposition of the likelihood ratio test for sphericity into two independent tests and a suitably developed decomposition of the characteristic function of the logarithm of the likelihood ratio test statistic to test independence in a set of variates, we may obtain extremely well-fitting near-exact distributions for both test statistics. Since both test statistics have the distribution of the product of independent Beta random variables, it is possible to obtain near-exact distributions for both statistics in the form of Generalized Near-Integer Gamma distributions or mixtures of these distributions. For the independence test statistic, numerical studies and comparisons with asymptotic distributions proposed by other authors show the extremely high accuracy of the near-exact distributions developed as approximations to the exact distribution. Concerning the sphericity test statistic, comparisons with formerly developed near-exact distributions show the advantages of these new near-exact distributions.

1. Introduction

Let $\mathbf{X}$ be a $p \times 1$ vector with a p-multivariate Normal distribution with expected value $\boldsymbol{\mu}$ and variance–covariance matrix Σ, denoted by

$$\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma). \tag{1}$$

Then, the $[2/(n+1)]$-th power of the likelihood ratio test statistic to test the null hypothesis of independence,

$$H_{01} : \Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2) \tag{2}$$

based on a sample of size $n+1$, is the statistic

$$\Lambda_1 = \frac{|V|}{\prod_{j=1}^{p} V_j} \tag{3}$$

where the $p \times p$ matrix V is either the MLE (Maximum Likelihood Estimator) of Σ, the sample matrix of sum of squares and products of deviations from the sample mean or the sample variance–covariance matrix of the p variables in $\mathbf{X}$ and V_j is the j-th diagonal element of V. The statistic in (3) is a particular case of the generalized Wilks Λ statistic used to test the independence of p groups of variables, when each of the groups contains only one variable (see [1], [2, Chap. 9]). Near-exact distributions for this statistic are thus readily available from the results in [3–5].

* Corresponding address: Mathematics Department, Faculty of Sciences and Technology, The New University of Lisbon, Quinta da Torre, 2829-516 Caparica, Portugal.
E-mail addresses: cmac@fct.unl.pt (C.A. Coelho), fjm@fct.unl.pt (F.J. Marques).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2009.09.012

In a simple way, we may say that near-exact distributions are asymptotic distributions built using an entirely different concept. These distributions are built in such a way that the major part of the exact c.f. (characteristic function) of the statistic is left unchanged and the remaining part is replaced by an asymptotic function, so that:

(i) if we denote by $\Phi^*(t)$ the part of the exact c.f. of the statistic that is replaced by $\Phi^*_n(t)$, where, for simplicity of notation, n is used to denote any and every parameter in the distribution of that statistic, we have

$$\lim_{n\to\infty}\int_{-\infty}^{+\infty}\left|\frac{\Phi^*_n(t)-\Phi^*(t)}{t}\right|\mathrm{d}t=0 \Longleftrightarrow \lim_{n\to\infty}\Phi^*_n(t)=\Phi^*(t),$$

with this replacement yielding what we call the near-exact c.f., in such a way that,

(ii) the near-exact distribution, obtained by inversion of the near-exact c.f., corresponds to a known and manageable distribution, from which the computation of p-values and quantiles is rendered easy.

However, given the specificity of the case under consideration, some further development may be sought. In particular, it is desirable to obtain a simpler formulation for the shape parameters of the Gamma distributions involved in the part of the distribution of Λ_1 left untouched. In addition, it is to be hoped that these parameters could be obtained with less effort than in the general case. These details will be addressed in Section 2.

On the other hand, the $[2/(n+1)]$-th power of the likelihood ratio test statistic to test the sphericity hypothesis on Σ, based on a sample of size $n+1$, that is, to test the null hypothesis

$$H_0 : \Sigma = \sigma^2 I_p \quad (\sigma^2 \text{ unspecified}) \tag{4}$$

is the statistic (see [2, Chap. 10])

$$\Lambda = p^p\,\frac{|V|}{(\mathrm{tr}V)^p}. \tag{5}$$

Well-fitting near-exact distributions have already been developed for this statistic by Marques and Coelho in [6]. In this paper we will show that even better near-exact distributions may be obtained for this statistic by taking as a basis the near-exact distributions developed for the statistic Λ_1 in (3) and the decomposition performed on its characteristic function. These near-exact distributions for Λ will be obtained from a decomposition of the statistic Λ in (5), which may be written as

$$\Lambda = \Lambda_1\,\Lambda_2, \tag{6}$$

where Λ_1 is the statistic in (3) and

$$\Lambda_2 = p^p\,\frac{\prod_{j=1}^{p} V_j}{(\mathrm{tr}V)^p}, \tag{7}$$

is the $[2/(n+1)]$-th power of the likelihood ratio test statistic to test the hypothesis

$$H_{02|01} : \sigma_1^2=\sigma_2^2=\cdots=\sigma_p^2 \quad \text{(given that, or, assuming that the } p \text{ variables in } \mathbf{X} \text{ are independent)} \tag{8}$$

based on p independent estimates of the variances of the variables in $\mathbf{X}$, one for each σ_j^2, from samples of size $n+1$.

The statistic in (7) may be derived from the likelihood ratio test statistic for the equality of p variance–covariance matrices (see Ch. 10 in [2]), taking each matrix to have dimensions 1×1 (the p groups consist of one variable each).

We may write for H_0 in (4), H_{01} in (2) and H_{02} in (8),

$$H_0 = H_{02|01} \circ H_{01},$$

to be read as "$H_{02|01}$ after H_{01}", meaning that we may test H_0 in two steps: (i) testing first H_{01}, that is, if the p variables in (1) are independent and (ii) once the hypothesis of independence of the p variables is not rejected, testing then if they all have the same variance. Under H_0 in (4) the two test statistics Λ_1 and Λ_2 in (6) are independent (see Ch. 10, subsec. 10.7.3 in [2]). This way to look at this test will enable us to obtain even better near-exact distributions than the already available ones in [6].

As a side note we may observe that the test statistic in (7) may be used, under a slightly different setting, to test the null hypothesis of equality of variances in (8), without any conditioning as long as the p estimators V_j are based on p independent samples, in which case those samples may have different sizes.

We should note that the requirement that the p estimators V_j in (7) have to be independent is indeed met even if the p estimators V_j come from a multivariate sample of size $n+1$ of the p variables in $\mathbf{X}$, once the null hypothesis of independence of the p variables is not rejected, since then the matrix V will have a Wishart distribution with n degrees of freedom and parameter matrix Σ, the matrix in (2), so that the diagonal elements of V are independent (through Theorem 7.3.5 in [2]).

2. Near-exact distributions for the likelihood ratio test statistic of independence

In order to obtain the c.f. of $W_1 = -\log\Lambda_1$ we may consider Theorems 9.3.2 and 9.3.3 of [2] which state that, for a sample of size $n+1$, Λ_1 in (3) has the same distribution as $\prod_{j=1}^{p-1} Y_j$, with

$$Y_j \sim B\left(\frac{n-p+j}{2},\frac{p-j}{2}\right) \quad (j=1,\dots,p-1) \tag{9}$$

where, under H_{01} in (2), the $p-1$ random variables Y_j in (9) are independent. Then, since we know that

$$E\left(Y_j^h\right)=\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-p+j}{2}+h\right)}{\Gamma\left(\frac{n}{2}+h\right)\Gamma\left(\frac{n-p+j}{2}\right)},\quad\left(h>-\frac{n-p+j}{2}\right)$$

we have, for $i=(-1)^{1/2}$,

$$\Phi_{W_1}(t)=E\left(\mathrm{e}^{\mathrm{i}tW_1}\right)=\prod_{j=1}^{p-1}E\left(Y_j^{-\mathrm{i}t}\right)=\prod_{j=1}^{p-1}\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-p+j}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)\Gamma\left(\frac{n-p+j}{2}\right)}. \tag{10}$$

Now, in order to be able to obtain a suitable decomposition of the c.f. of W_1 we may either consider the results and developments in Section 5 of [3], taking $p_k=1$ for $k=1,\ldots,m$ and $k^*=\lfloor p/2\rfloor$, or we may take a different approach which will indeed enable us to obtain simpler expressions for the shape parameters of the Gamma distributions involved in the part of the distribution of W_1 which will be left unchanged. We will take this second approach.

The following Lemma gives the c.f. of W_1, for both even and odd p, in a form that is suitable for the development of near-exact distributions for both W_1 and Λ_1.

Lemma 1. *Under H_{01} in (2), taking $k^*=\lfloor p/2\rfloor$, the c.f. of $W_1=-\log\,\Lambda_1$ (where Λ_1 is the statistic in (3)), may be written in the form*

$$\Phi_{W_1}(t)=\Phi_1^*(t)\,\Phi_2^*(t), \tag{11}$$

where

$$\Phi_1^*(t)=\left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)}\right)^{k^*}$$

is the c.f. of the sum of k^* independent Logbeta r.v.'s with parameters $(n-1)/2$ and $1/2$ and

$$\Phi_2^*(t)=\prod_{k=1}^{p-2}\left(\frac{n-1-k}{2}\right)^{\left\lfloor\frac{p-k}{2}\right\rfloor}\left(\frac{n-1-k}{2}-\mathrm{i}t\right)^{-\left\lfloor\frac{p-k}{2}\right\rfloor}$$

is the c.f. of a GIG (Generalized Integer Gamma) distribution [1] of depth $p-2$ with rate parameters $(n-1-k)/2$ and shape parameters $\left\lfloor\frac{p-k}{2}\right\rfloor$ $(k=1,\ldots,p-2)$, that is the distribution of the sum of $p-2$ independent Gamma r.v.'s with the given rate and integer shape parameters.

Proof. See Appendix.

Then, taking into account the fact that a single Logbeta distribution may be represented in the form of an infinite mixture of Exponential distributions [7], a sum of independent Logbeta r.v.'s with either the same or different parameters may thus be represented in the form of an infinite mixture of sums of independent Exponentials, that is an infinite mixture of GIG distributions. Then, taking into account the fact that the GIG distribution itself may be seen as a mixture of Gamma distributions [8], the replacement of the sum of independent Logbeta r.v.'s by a single Gamma distribution or by a (finite) mixture of Gamma distributions seems to be most adequate.

Thus, near-exact distributions for W_1 may then be obtained in the form of a GNIG (Generalized Near-Integer Gamma) distribution [3] or mixtures of GNIG distributions by replacing $\Phi_1^*(t)$ by the c.f. of a Gamma distribution or the c.f. of a mixture of Gamma distributions. These near-exact distributions will match, by construction, the first two, four or six exact moments of W_1.

Theorem 2. *Using for $\Phi_1^*(t)$ in (11), the approximations:*

$$\sum_{k=1}^{h/2}\theta_k\,\mu^{\delta_k}(\mu-\mathrm{i}t)^{-\delta_k},\quad h=2,4,6, \tag{12}$$

for $\theta_k,\mu,\delta_k>0$ and $\sum_{k=1}^{h/2}\theta_k=1$, such that, for $j=1,\ldots,h$,

$$\left.\frac{\partial^j}{\partial t^j}\sum_{k=1}^{h/2}\theta_k\,\mu^{\delta_k}(\mu-\mathrm{i}t)^{-\delta_k}\right|_{t=0}=\left.\frac{\partial^j}{\partial t^j}\,\Phi_1^*(t)\right|_{t=0}, \tag{13}$$

we obtain as near-exact distributions for W_1, a GNIG distribution, for $h=2$, or, for $h=4$ or 6, a mixture of 2 or 3 GNIG distributions, with cdf's given by (using the notation in (19) of [6])

$$\sum_{k=1}^{h/2}\theta_kF(w|r_1,\ldots,r_{p-2},\delta_k;\lambda_1,\ldots,\lambda_{p-2},\mu) \tag{14}$$

where

$$r_j = \left\lfloor \frac{p-j}{2} \right\rfloor, \qquad \lambda_j = \frac{n-1-j}{2}, \quad (j = 1, \ldots, p-2), \tag{15}$$

μ and $\delta_1, \ldots, \delta_{h/2}$ and $\theta_1, \ldots, \theta_{h/2-1}$, are obtained from the numerical solution of the system of h equations

$$\sum_{k=1}^{h/2} \theta_k \frac{\Gamma(\delta_k + j)}{\Gamma(\delta_k)} \mu^{-j} = \mathrm{i}^{-j} \left. \frac{\partial^j}{\partial t^j} \Phi_1^*(t) \right|_{t=0} \quad (j = 1, \ldots, h) \tag{16}$$

for these parameters, with $\theta_{h/2} = 1 - \sum_{k=1}^{h/2-1} \theta_k$.

Proof. If in the characteristic function of W_1 in (11) we replace $\Phi_1^*(t)$ by $\sum_{k=1}^{h/2} \theta_k \, \mu^{\delta_k} (\mu - \mathrm{i}t)^{-\delta_k}$ we obtain

$$\Phi_{W_1}(t) \approx \sum_{k=1}^{h/2} \theta_k \, \mu^{\delta_k} (\mu - \mathrm{i}t)^{-\delta_k} \prod_{j=1}^{p-2} \left(\frac{n-1-j}{2} \right)^{\left\lfloor \frac{p-j}{2} \right\rfloor} \left(\frac{n-1-j}{2} - \mathrm{i}t \right)^{-\left\lfloor \frac{p-j}{2} \right\rfloor},$$

that is the characteristic function of a GNIG distribution, for $h = 2$, or of a mixture of 2 or 3 GNIG distributions, for $h = 4$ or 6, with cdf's given by (14), with each component in the mixture being clearly the sum of $p-1$ independent Gamma random variables, $p-2$ of which with integer shape parameters r_j and rate parameters λ_j given by (15), and a further Gamma random variable with rate parameter $\delta_k > 0$ and shape parameter μ. The parameters δ_k, μ and θ_k are determined in such a way that (16) holds (for which indeed there is a simple analytical solution for $h = 2$). □

Corollary 3. *Distributions with cdf's given by*

$$1 - \sum_{k=1}^{h/2} \theta_k F(-\log z | r_1, \ldots, r_{p-2}, \delta_k; \lambda_1, \ldots, \lambda_{p-2}, \mu), \quad (h = 2, 4, 6)$$

where the parameters are the same as in Theorem 2, and $0 < z < 1$ represents the running value of the statistic $\Lambda_1 = \mathrm{e}^{-W_1}$, may be used as near-exact distributions for this statistic.

Proof. Since the near-exact distributions developed in Theorem 2 were for the random variable $W_1 = -\log \Lambda_1$ we only need to mind the relation

$$F_{\Lambda_1}(z) = 1 - F_{W_1}(-\log z)$$

where $F_{\Lambda_1}(\cdot)$ is the cumulative distribution function of Λ_1 and $F_{W_1}(\cdot)$ is the cumulative distribution function of W_1, in order to obtain the corresponding near-exact distributions for Λ_1. □

Indeed, in order to obtain near-exact α-quantiles for Λ_1, we do not even need the near-exact distributions for Λ_1, since if we consider the relation

$$\Lambda_1(\alpha) = \mathrm{e}^{-W_1(1-\alpha)},$$

where $\Lambda_1(\alpha)$ is the α-quantile of Λ_1 and $W_1(1-\alpha)$ is the $(1-\alpha)$-quantile of W_1 we may easily obtain the near-exact α-quantiles of Λ_1 from the corresponding $(1-\alpha)$-quantiles of W_1.

3. Near-exact distributions for the likelihood ratio test statistic of sphericity

Lemma 4. *The c.f. of $W = -\log \Lambda$, where Λ is the test statistic in (5) may be written, for $k^* = \lfloor p/2 \rfloor$, as*

$$\Phi_W(t) = \Phi_1^{**}(t) \, \Phi_2^{**}(t) \tag{17}$$

where

$$\Phi_1^{**}(t) = \prod_{j=p-k^*+1}^{p} \frac{\Gamma\left(\frac{n}{2} + \frac{j-1}{p}\right) \Gamma\left(\frac{n+1}{2} - \mathrm{i}t\right)}{\Gamma\left(\frac{n+1}{2}\right) \Gamma\left(\frac{n}{2} + \frac{j-1}{p} - \mathrm{i}t\right)} \prod_{j=1}^{p-k^*} \frac{\Gamma\left(\frac{n}{2} + \frac{j-1}{p}\right) \Gamma\left(\frac{n}{2} - \mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n}{2} + \frac{j-1}{p} - \mathrm{i}t\right)}$$

is the c.f. of the sum of p independent Logbeta r.v.'s, k^* of which with parameters $(n+1)/2$ and $(j-1)/p - 1/2$ ($j = p - k^* + 1, \ldots, p$) and the remaining $p - k^*$ with parameters $n/2$ and $(j-1)/p$ ($j = 1, \ldots, p - k^*$) and

$$\Phi_2^{**}(t) = \prod_{k=1}^{p-1} \left(\frac{n-k}{2} \right)^{\left\lfloor \frac{p-k+1}{2} \right\rfloor} \left(\frac{n-k}{2} - \mathrm{i}t \right)^{-\left\lfloor \frac{p-k+1}{2} \right\rfloor}$$

is the c.f. of a GIG distribution of depth $p-1$, with rate parameters $(n-k)/2$ and shape parameters $\left\lfloor \frac{p-k+1}{2} \right\rfloor$ ($k = 1, \ldots, p-1$).

Proof. See Appendix.

Then, by replacing $\Phi_1^{**}(t)$ in (17) by the c.f. of a Gamma distribution or the c.f. of a mixture of Gamma distributions, we will get near-exact distributions for $W = -\log \Lambda$ in the form of a GNIG distribution or mixtures of GNIG distributions.

As we will see in the next section, these near-exact distributions provide better approximations than the ones already available in [6].

Theorem 5. *Using for $\Phi_1^{**}(t)$ in (17), approximations similar in formulation to the ones used in Theorem 2 for $\Phi_1^{*}(t)$, with $\Phi_1^{*}(t)$ replaced by $\Phi_1^{**}(t)$ in (13), we obtain as near-exact distributions for $W = -\log \Lambda$, respectively a GNIG distribution, or a mixture of two or three GNIG distributions, all of depth p, which respectively match the first two, four or six exact moments and which have cdf's (using the notation in (19) of [6])*

$$\sum_{k=1}^{h/2} \theta_k F(w|r_1,\ldots,r_{p-1},\delta_k;\lambda_1,\ldots,\lambda_{p-1},\mu), \tag{18}$$

where

$$r_j = \left\lfloor \frac{p-j+1}{2} \right\rfloor, \qquad \lambda_j = \frac{n-j}{2}, \quad (j=1,\ldots,p-1), \tag{19}$$

$\theta_k (k = 1,\ldots,h/2-1)$, μ, $\delta_k (k=1,\ldots,h/2)$, in (18), with $\theta_{h/2} = 1-\sum_{k=1}^{h/2-1}\theta_k$, are obtained from the numerical solution of a system of h equations similar to the ones in (16) with $\Phi_1^{*}(t)$ replaced by $\Phi_1^{**}(t)$.

Proof. Similar to the proof of Theorem 2. For details see Appendix.

Corollary 6. *Distributions with cdf's given by*

$$1-\sum_{k=1}^{h/2} \theta_k F(-\log z|r_1,\ldots,r_{p-1},\delta_k;\lambda_1,\ldots,\lambda_{p-1},\mu), \quad (h=2,4,6)$$

where the parameters are the same as in Theorem 5, and $0 < z < 1$ represents the running value of the statistic $\Lambda = \mathrm{e}^{-W}$, may be used as near-exact distributions for this statistic.

The proof of this corollary is in all respects similar to the proof of Corollary 3 and also similar considerations to the ones right after Corollary 3, concerning the computation of near-exact quantiles of the statistics W_1 and Λ_1, apply here to the computation of near-exact quantiles of the statistics W and Λ.

4. Numerical and comparative studies

In order to assess the proximity between the near-exact approximations developed and the exact distribution, we will use two measures of proximity,

$$\Delta_1 = \int_{-\infty}^{\infty} |\phi_W(t)-\phi_n(t)|\,\mathrm{d}t \quad \text{and} \quad \Delta_2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\frac{\phi_W(t)-\phi_n(t)}{t}\right| \mathrm{d}t, \tag{20}$$

with

$$\max_{w\in S} |f_W(w)-f_n(w)| \le \frac{1}{2\pi}\Delta_1 \quad \text{and} \quad \max_{w\in S} |F_W(w)-F_n(w)| \le \Delta_2, \tag{21}$$

where W represents a continuous random variable defined on S with cdf $F_W(w)$, density function $f_W(w)$ and characteristic function $\phi_W(t)$, and $\phi_n(t)$, $F_n(y)$ and $f_n(y)$ represent respectively the near-exact characteristic, distribution and density function of W. For further details on these measures see [6].

In this section we will denote by GNIG, M2GNIG and M3GNIG respectively the near-exact distributions based on a GNIG or a mixture of two or three GNIG distributions, for both the likelihood ratio statistics of independence and sphericity. In Section 4.2 we will use GNIGpre, M2GNIGpre and M3GNIGpre to denote the previous near-exact distributions based respectively on a GNIG distribution or on a mixture of two or three of such distributions, obtained in [6].

4.1. Studies for the independence test statistic

Mudholkar et al. in [9] developed a Normal approximation to the distribution of the likelihood ratio test statistic used for testing H_{01} in (2). These authors presented numerical studies comparing their Normal approximation with the approximations due to Box and Bartlett [10,11].

Since the asymptotic Normal approximation from Mudholkar et al. in [9] yields indeed for $\log \Lambda_1$ a non-central generalized Gamma distribution, whose c.f. is not manageable, in order to compare the performance of the near-exact distributions developed with this Normal asymptotic approximation, instead of using measures Δ_1 and Δ_2, we decided to use a similar method to the one used in [9] to assess the performance of their Normal asymptotic approximation.

Table 1
Values of the tail probability error $=$ (*approx.prob*-α) for the Mudholkar et al. [9] asymptotic distribution (MTL) and the near-exact distributions GNIG and M2GNIG, for samples of size $n+1$.

p	n		α					
			0.005	0.01	0.05	0.10	0.20	0.50
3	6	MTL	1.90×10^{-4}	1.10×10^{-4}	-8.60×10^{-4}	-1.64×10^{-3}	-1.71×10^{-3}	2.81×10^{-3}
		GNIG	1.29×10^{-5}	2.72×10^{-5}	-1.23×10^{-5}	-7.92×10^{-5}	-1.65×10^{-4}	-5.75×10^{-5}
		M2GNIG	-9.91×10^{-6}	-9.77×10^{-7}	-4.20×10^{-6}	-1.82×10^{-6}	5.42×10^{-6}	-6.78×10^{-6}
	13	MTL	1.80×10^{-4}	1.10×10^{-4}	-9.00×10^{-4}	-1.67×10^{-3}	-1.73×10^{-3}	2.79×10^{-3}
		GNIG	1.26×10^{-5}	3.47×10^{-6}	-2.61×10^{-6}	-1.76×10^{-5}	-3.30×10^{-5}	-5.33×10^{-6}
		M2GNIG	7.79×10^{-6}	-1.85×10^{-6}	1.75×10^{-6}	4.88×10^{-7}	4.69×10^{-7}	-1.13×10^{-6}
6	9	MTL	5.00×10^{-5}	6.00×10^{-5}	1.00×10^{-5}	-1.40×10^{-4}	-3.70×10^{-4}	-2.80×10^{-4}
		GNIG	2.05×10^{-6}	2.97×10^{-6}	4.25×10^{-6}	9.65×10^{-7}	-7.81×10^{-6}	-1.83×10^{-5}
		M2GNIG	1.58×10^{-8}	-8.79×10^{-8}	1.85×10^{-8}	-3.24×10^{-8}	-3.26×10^{-8}	7.92×10^{-8}
	16	MTL	8.00×10^{-5}	7.00×10^{-5}	-1.60×10^{-4}	-4.10×10^{-4}	-6.00×10^{-4}	4.00×10^{-5}
		GNIG	1.15×10^{-6}	1.59×10^{-6}	1.43×10^{-6}	4.31×10^{-7}	4.27×10^{-6}	-6.82×10^{-6}
		M2GNIG	2.80×10^{-8}	4.49×10^{-8}	-1.08×10^{-8}	4.80×10^{-9}	2.21×10^{-8}	1.56×10^{-8}
10	13	MTL	-2.00×10^{-5}	-1.00×10^{-5}	1.20×10^{-4}	1.60×10^{-4}	6.00×10^{-5}	-4.70×10^{-4}
		GNIG	2.12×10^{-7}	3.25×10^{-7}	5.45×10^{-7}	3.26×10^{-7}	-4.52×10^{-7}	-1.87×10^{-6}
		M2GNIG	7.02×10^{-11}	-2.87×10^{-12}	-7.41×10^{-10}	-9.29×10^{-10}	-5.28×10^{-11}	2.01×10^{-9}
	20	MTL	2.00×10^{-5}	2.00×10^{-5}	-6.00×10^{-5}	-1.60×10^{-4}	-2.80×10^{-4}	-3.30×10^{-4}
		GNIG	1.97×10^{-7}	2.82×10^{-7}	3.70×10^{-7}	1.32×10^{-7}	-4.80×10^{-7}	-1.25×10^{-6}
		M2GNIG	1.50×10^{-10}	-2.32×10^{-10}	-5.39×10^{-10}	-6.23×10^{-10}	2.71×10^{-10}	1.48×10^{-9}

Table 2
Values of the measures Δ_1 and Δ_2 for the near-exact distributions, for samples of size $n+1$.

p	n	Δ_1			Δ_2		
		GNIG	M2GNIG	M3GNIG	GNIG	M2GNIG	M3GNIG
3	6	5.8×10^{-2}	5.7×10^{-3}	2.4×10^{-3}	5.1×10^{-4}	2.1×10^{-5}	4.6×10^{-6}
	13	2.7×10^{-2}	1.9×10^{-3}	2.7×10^{-5}	9.4×10^{-5}	3.1×10^{-6}	3.3×10^{-8}
6	9	3.3×10^{-4}	3.5×10^{-6}	3.6×10^{-8}	2.4×10^{-5}	1.7×10^{-7}	1.3×10^{-9}
	16	2.8×10^{-4}	2.4×10^{-6}	2.1×10^{-9}	9.5×10^{-6}	5.6×10^{-8}	3.5×10^{-11}
10	13	1.9×10^{-5}	4.2×10^{-8}	6.2×10^{-12}	2.0×10^{-6}	3.2×10^{-9}	3.5×10^{-13}
	20	2.4×10^{-5}	5.1×10^{-8}	4.6×10^{-11}	1.4×10^{-6}	2.1×10^{-9}	1.5×10^{-12}
20	23	5.2×10^{-7}	1.4×10^{-10}	3.0×10^{-14}	7.8×10^{-8}	1.5×10^{-11}	2.7×10^{-15}
	50	1.1×10^{-6}	2.8×10^{-10}	7.4×10^{-14}	4.9×10^{-8}	9.7×10^{-12}	2.1×10^{-15}
	100	7.4×10^{-7}	1.2×10^{-10}	2.5×10^{-14}	1.5×10^{-8}	1.8×10^{-12}	3.2×10^{-16}
50	53	5.8×10^{-9}	1.0×10^{-13}	1.5×10^{-19}	1.1×10^{-9}	1.5×10^{-14}	1.9×10^{-20}
	100	1.1×10^{-6}	2.8×10^{-10}	1.9×10^{-17}	4.9×10^{-8}	9.7×10^{-12}	7.5×10^{-19}
	150	2.3×10^{-8}	4.6×10^{-13}	4.3×10^{-19}	8.5×10^{-10}	1.3×10^{-14}	1.0×10^{-20}
	200	2.0×10^{-8}	3.2×10^{-13}	3.3×10^{-19}	5.3×10^{-10}	6.5×10^{-15}	5.7×10^{-21}
	500	9.8×10^{-9}	7.1×10^{-14}	5.0×10^{-20}	1.0×10^{-12}	5.4×10^{-16}	3.2×10^{-22}

We used the exact quantiles for Λ_1 computed directly from the numerical inversion of the c.f. of $\log \Lambda_1$ by using the Gil-Pelaez inversion formulas (see [12]) which gives us a precision at least equal to the one used in [9] in terms of exact quantiles, which in turn give, for the Normal asymptotic approximation of [9], exactly the same results obtained by these authors.

However, given that the exact quantiles computed in this way have a precision that does not go beyond 12 digits, and given that this precision is not enough for making comparisons with the near-exact distribution M3GNIG, which requires a higher precision, we have used in Table 1 only the near-exact distributions GNIG and M2GNIG.

In Table 1 the errors displayed are evaluated using the exact same method used by Mudholkar et al. in [9], the difference between the approximate and the exact tail probabilities. The cases considered are a subset of the ones considered by Mudholkar et al. in [9]. We may observe that the errors obtained when using the near-exact distributions are always much smaller than the ones obtained for the Normal approximation of Mudholkar et al. in [9], mainly for larger values of p.

In Table 2 we use measures Δ_1 and Δ_2 to better assess the relative performance of the three near-exact distributions GNIG, M2GNIG and M3GNIG as approximating distributions for the independence test statistic.

From Table 2 we may easily see that the near-exact distribution M3GNIG provides a better approximation than the other two near-exact distributions and we may also see that the near-exact distribution M2GNIG always outperforms the GNIG near-exact distribution. The values exhibited for the M3GNIG distribution for both measures, mainly for the measure Δ_2, which represents an upper bound for the absolute value of the difference between its cdf and the exact cdf, would lead us to recommend its use as a replacement for the exact distribution, mainly for larger values of p. The three near-exact distributions display an asymptotic behavior, that is, smaller values for the two measures, both for increasing sample sizes and increasing number of variables, although for larger values of p we need to consider large enough sample sizes in order to be able to observe their asymptotic behavior in terms of sample size.

Table 3
Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log \Lambda$, for $p = 4$, $n = 6$ and $p = 5$, $n = 7$, for samples of size $n + 1$.

	$p = 4, n = 6$		$p = 5, n = 7$	
	Δ_1	Δ_2	Δ_1	Δ_2
GNIG	1.198×10^{-4}	7.564×10^{-6}	8.803×10^{-5}	7.226×10^{-6}
M2GNIG	1.765×10^{-6}	6.906×10^{-8}	8.104×10^{-7}	4.330×10^{-8}
M3GNIG	3.706×10^{-8}	3.706×10^{-10}	1.551×10^{-8}	6.259×10^{-10}

Table 4
Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log \Lambda$, for $p = 7$, $n = 9$ and $p = 10$, $n = 12$, for samples of size $n + 1$.

	$p = 7, n = 9$		$p = 10, n = 12$	
	Δ_1	Δ_2	Δ_1	Δ_2
GNIG	1.242×10^{-5}	1.330×10^{-6}	1.058×10^{-6}	1.383×10^{-7}
M2GNIG	4.281×10^{-8}	3.163×10^{-9}	8.994×10^{-10}	8.370×10^{-11}
M3GNIG	1.765×10^{-10}	1.018×10^{-11}	2.779×10^{-13}	2.048×10^{-14}

Table 5
Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log \Lambda$, for $p = 4, 5, 7$ and $n = 50$, for samples of size $n + 1$.

	$p = 4, n = 50$		$p = 5, n = 50$		$p = 7, n = 50$	
	Δ_1	Δ_2	Δ_1	Δ_2	Δ_1	Δ_2
GNIG	3.048×10^{-5}	1.887×10^{-7}	2.981×10^{-5}	2.668×10^{-7}	7.446×10^{-6}	1.044×10^{-7}
M2GNIG	8.425×10^{-8}	3.472×10^{-10}	9.222×10^{-8}	5.765×10^{-10}	1.123×10^{-8}	1.137×10^{-10}
M3GNIG	3.872×10^{-10}	1.199×10^{-12}	1.550×10^{-10}	7.504×10^{-13}	8.671×10^{-12}	7.047×10^{-14}

Table 6
Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log \Lambda$, for $p = 10, 20, 30$ and $n = 50$, for samples of size $n + 1$.

	$p = 10, n = 50$		$p = 20, n = 50$		$p = 30, n = 50$	
	Δ_1	Δ_2	Δ_1	Δ_2	Δ_1	Δ_2
GNIG	1.204×10^{-6}	2.561×10^{-8}	1.001×10^{-7}	4.755×10^{-9}	1.816×10^{-8}	1.460×10^{-9}
M2GNIG	5.382×10^{-10}	8.438×10^{-12}	1.138×10^{-11}	4.030×10^{-13}	8.197×10^{-13}	4.922×10^{-14}
M3GNIG	1.959×10^{-13}	2.515×10^{-15}	1.131×10^{-15}	3.321×10^{-17}	3.521×10^{-17}	1.755×10^{-18}

4.2. Studies for the sphericity test statistic

Tables 3–7 present the values of the measures Δ_1 and Δ_2 given in (20) for the new near-exact distributions, GNIG, M2GNIG and M3GNIG, developed in this paper for the likelihood ratio test statistic used for testing sphericity. In this sub-section our purpose is to assess the quality of the new near-exact distributions and compare them with the ones already developed, GNIGpre, M2GNIGpre and M3GNIGpre, using a different method, in [6]. In order to achieve our purpose we have considered the exact same values for n and p already considered in the numerical studies presented in that reference.

Comparing Tables 3 and 4 with Tables 1 and 2 in [6] we may observe that the values for the new approximations are always better with the exception of M3GNIG for $p = 5$, $n = 7$ and $p = 7$, $n = 9$.

Comparing Tables 5 and 6 with Tables 3 and 4 in [6] we may verify that in most cases we have for the new near-exact approximations smaller values for the measures Δ_1 and Δ_2, with the exception of the near-exact distributions M2GNIG and M3GNIG for the cases in Table 5. The new near-exact approximations developed in this paper also exhibit the good asymptotic properties exhibited by the near-exact approximations in [6].

We may say, as a general conclusion, that the new near-exact distributions perform better than the ones developed in [6] for large values of n with p large enough ($p \geq 10$). Moreover in Table 7 we may see that for large values of p and values of n close to p we also have better values of both measures for the near-exact distributions developed in this paper.

5. Conclusions

In this paper we have shown how, using a decomposition of the null hypothesis into two independent hypotheses, we may use the induced factorization of the overall test statistic to obtain very accurate near-exact distributions for the sphericity test statistic. These new near-exact distributions developed for the sphericity test statistic are more accurate than the ones developed in [6] for larger values of p ($p \geq 10$). As a by-product we have also obtained near-exact distributions for the independence test statistic which show a much better precision than the Normal approximation in [9].

Moreover, the process used to factorize the characteristic functions involved allowed us to obtain near-exact distributions almost simultaneously for the independence and the sphericity test statistics, and also to obtain simple expressions for the shape parameters of $\Phi_2^*(t)$ in (11) and $\Phi_2^{**}(t)$ in (17), with the shape parameters for the near-exact distributions for the sphericity test statistic bearing much simpler expressions than the ones for the near-exact distributions in [6].

Table 7
Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log \Lambda$, for $p = 10, 20, 30$ and $n = 12, 22, 32$, for samples of size $n+1$.

	$p=10, n=12$		$p=20, n=22$		$p=30, n=32$	
	Δ_1	Δ_2	Δ_1	Δ_2	Δ_1	Δ_2
GNIG	1.058×10^{-6}	1.383×10^{-7}	3.526×10^{-8}	6.034×10^{-9}	5.043×10^{-9}	9.685×10^{-10}
GNIGpre	8.940×10^{-6}	1.171×10^{-6}	3.634×10^{-7}	6.221×10^{-8}	5.178×10^{-8}	9.945×10^{-9}
M2GNIG	8.994×10^{-10}	8.380×10^{-11}	3.664×10^{-12}	4.587×10^{-13}	1.543×10^{-13}	2.185×10^{-14}
M2GNIGpre	3.394×10^{-9}	3.189×10^{-10}	1.173×10^{-11}	1.470×10^{-12}	4.525×10^{-13}	6.408×10^{-14}
M3GNIG	2.779×10^{-13}	2.048×10^{-14}	2.956×10^{-16}	3.014×10^{-17}	4.413×10^{-18}	5.122×10^{-19}
M3GNIGpre	3.601×10^{-12}	2.706×10^{-13}	1.104×10^{-15}	1.128×10^{-16}	1.766×10^{-17}	2.051×10^{-18}

Acknowledgment

This research was financially supported by the Portuguese Foundation for Science and Technology, through the Center for Mathematics and its Applications (CMA) from the New University of Lisbon. The authors would also like to express their appreciation for the insightful comments and suggestions made by the Reviewers and the assigned Editor.

Appendix. Proofs of Lemmas 1 and 4 and Theorem 5

Proof of Lemma 1. We will need to consider separately the two cases of even and odd p. For even p, from (10) we may write

$$\Phi_{W_1}(t) = \Phi_1(t)\,\Phi_2(t),$$

with

$$\Phi_1(t) = \prod_{\substack{j=1\\ step2}}^{p-1} \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-p+j}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)\Gamma\left(\frac{n-p+j}{2}\right)}, \qquad \Phi_2(t) = \prod_{\substack{j=2\\ step2}}^{p-2} \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-p+j}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)\Gamma\left(\frac{n-p+j}{2}\right)},$$

where, for even j, $(p-j)/2 \in \mathbb{N}$. Thus, using for $z \in \mathbb{C} \setminus \{0, -1, -2, \ldots\}$ and $n \in \mathbb{N}$,

$$\frac{\Gamma(z+n)}{\Gamma(z)} = \prod_{k=0}^{n-1}(z+k),$$

we may write

$$\begin{aligned}
\Phi_2(t) &= \prod_{\substack{j=2\\ step2}}^{p-2} \prod_{k=0}^{\frac{p-j}{2}-1} \left(\frac{n-p+j}{2}+k\right)\left(\frac{n-p+j}{2}+k-\mathrm{i}t\right)^{-1} \\
&= \prod_{\substack{j=2\\ step2}}^{p-2} \prod_{k=0}^{\frac{p-j}{2}-1} \left(\frac{n-p+j}{2}+\frac{p-j}{2}-1-k\right)\left(\frac{n-p+j}{2}+\frac{p-j}{2}-1-k-\mathrm{i}t\right)^{-1} \\
&= \prod_{\substack{j=2\\ step2}}^{p-2} \prod_{k=1}^{\frac{p-j}{2}} \left(\frac{n}{2}-k\right)\left(\frac{n}{2}-k-\mathrm{i}t\right)^{-1} = \prod_{k=1}^{\frac{p-2}{2}} \left(\frac{n}{2}-k\right)^{\frac{p}{2}-k}\left(\frac{n}{2}-k-\mathrm{i}t\right)^{-\left(\frac{p}{2}-k\right)}
\end{aligned}$$

and

$$\begin{aligned}
\Phi_1(t) &= \prod_{\substack{j=1\\ step2}}^{p-1} \frac{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n-p+j}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)\Gamma\left(\frac{n-p+j}{2}\right)} \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)} \\
&= \left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)}\right)^{p/2} \prod_{\substack{j=1\\ step2}}^{p-1} \prod_{k=0}^{\frac{p-j-1}{2}-1} \left(\frac{n-p+j}{2}+k\right)\left(\frac{n-p+j}{2}+k-\mathrm{i}t\right)^{-1} \\
&= \left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)}\right)^{p/2} \prod_{\substack{j=1\\ step2}}^{p-1} \prod_{k=0}^{\frac{p-j-1}{2}-1} \left(\frac{n-p+j}{2}+\frac{p-j-1}{2}-1-k\right)
\end{aligned}$$

$$\times\left(\frac{n-p+j}{2}+\frac{p-j-1}{2}-1-k-\mathrm{i}t\right)^{-1}$$

$$=\left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)}\right)^{p/2}\prod_{k=1}^{\frac{p-2}{2}}\left(\frac{n-1}{2}-k\right)^{\frac{p}{2}-k}\left(\frac{n-1}{2}-k-\mathrm{i}t\right)^{-\left(\frac{p}{2}-k\right)},$$

so that we may finally write, for even p,

$$\Phi_{W_1}(t)=\left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)}\right)^{p/2}\prod_{k=1}^{p-2}\left(\frac{n-1-k}{2}\right)^{\frac{p}{2}-\lfloor\frac{k+1}{2}\rfloor}\left(\frac{n-1-k}{2}-\mathrm{i}t\right)^{-\frac{p}{2}+\lfloor\frac{k+1}{2}\rfloor}.$$

For odd p, we may write

$$\Phi_{W_1}(t)=\Phi_3(t)\,\Phi_4(t),$$

with

$$\Phi_3(t)=\prod_{\substack{j=1\\ step2}}^{p-2}\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-p+j}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)\Gamma\left(\frac{n-p+j}{2}\right)},\qquad \Phi_4(t)=\prod_{\substack{j=2\\ step2}}^{p-1}\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-p+j}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)\Gamma\left(\frac{n-p+j}{2}\right)},$$

where now it is for odd j that $(p-j)/2\in\mathbb{N}$, so that following similar steps to the ones used above to handle $\Phi_2(t)$, we may write

$$\Phi_3(t)=\prod_{\substack{j=1\\ step2}}^{p-2}\prod_{k=0}^{\frac{p-j}{2}-1}\left(\frac{n-p+j}{2}+k\right)\left(\frac{n-p+j}{2}+k-\mathrm{i}t\right)^{-1}$$

$$=\prod_{\substack{j=1\\ step2}}^{p-2}\prod_{k=1}^{\frac{p-j}{2}}\left(\frac{n}{2}-k\right)\left(\frac{n}{2}-k-\mathrm{i}t\right)^{-1}$$

$$=\prod_{k=1}^{\frac{p-1}{2}}\left(\frac{n}{2}-k\right)^{\frac{p-1}{2}-k}\left(\frac{n}{2}-k-\mathrm{i}t\right)^{-\left(\frac{p-1}{2}-k\right)}$$

and

$$\Phi_4(t)=\prod_{\substack{j=2\\ step2}}^{p-1}\frac{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n-p+j}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)\Gamma\left(\frac{n-p+j}{2}\right)}\,\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)}$$

$$=\left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)}\right)^{(p-1)/2}\prod_{\substack{j=2\\ step2}}^{p-1}\prod_{k=0}^{\frac{p-j-1}{2}-1}\left(\frac{n-p+j}{2}+k\right)\left(\frac{n-p+j}{2}+k-\mathrm{i}t\right)^{-1}$$

$$=\left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)}\right)^{(p-1)/2}\prod_{k=1}^{\frac{p-3}{2}}\left(\frac{n-1}{2}-k\right)^{\frac{p-1}{2}-k}\left(\frac{n-1}{2}-k-\mathrm{i}t\right)^{-\left(\frac{p-1}{2}-k\right)},$$

so that we may finally write, for odd p,

$$\Phi_{W_1}(t)=\left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-\mathrm{i}t\right)}\right)^{(p-1)/2}\prod_{k=1}^{p-2}\left(\frac{n-1-k}{2}\right)^{\frac{p-1}{2}-\lfloor\frac{k}{2}\rfloor}\left(\frac{n-1-k}{2}-t\right)^{-\frac{p-1}{2}+\lfloor\frac{k}{2}\rfloor}.$$

Then, taking $k^*=\lfloor p/2\rfloor$, we may write $\Phi_W(t)$ in the form (11), for any even or odd p. □

Proof of Lemma 4. The h-th null moment of the statistic in (7) is (see Chap. 10 in [2])

$$E\left(\Lambda_2^h\right)=p^{ph}\,\frac{\Gamma\left(\frac{np}{2}\right)}{\Gamma\left(\frac{np}{2}+hp\right)}\prod_{j=1}^{p}\frac{\Gamma\left(\frac{n}{2}+h\right)}{\Gamma\left(\frac{n}{2}\right)}$$

so that taking $W_2=-\log\,\Lambda_2$ and using the multiplication formula for the Gamma function

$$\Gamma(kz)=(2\pi)^{-(k-1)/2}\,k^{kz-1/2}\prod_{j=1}^{k}\Gamma\left(z+\frac{j-1}{k}\right),$$

we have

$$\begin{aligned}\Phi_{W_2}(t) &= E\left(e^{itW_2}\right) = E\left(e^{-it\log\Lambda_2}\right) = E\left(\Lambda_2^{-it}\right)\\ &= p^{-itp}\,\frac{\Gamma\left(\frac{np}{2}\right)}{\Gamma\left(\frac{np}{2}-itp\right)}\prod_{j=1}^{p}\frac{\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)}\\ &= \prod_{j=1}^{p}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}.\end{aligned} \tag{22}$$

Then, since

$$k^* = \left\lfloor \frac{p}{2}\right\rfloor = \begin{cases}\dfrac{p}{2} & \text{even } p\\ \dfrac{p-1}{2} & \text{odd } p,\end{cases}$$

for $p-k^*+1 \le j \le p$, we have

$$\frac{j-1}{p}+\frac{1}{2} \ge \frac{p-k^*}{p}+\frac{1}{2} = \begin{cases}\dfrac{p-p/2}{p}+\dfrac{1}{2}=1 & \text{even } p\\ \dfrac{p-(p-1)/2}{p}+\dfrac{1}{2}=1+\dfrac{1}{2p} & \text{odd p,}\end{cases}$$

and thus

$$\frac{j-1}{p}+\frac{1}{2} \ge 1.$$

Therefore, given that under H_0 in (4), Λ_1 and Λ_2 are independent, we may write, from (11) and (22), the c.f. of $W=-\log\,\Lambda$, where Λ is the statistic in (5), as

$$\begin{aligned}\Phi_W(t) &= \Phi_{W_1}(t)\,\Phi_{W_2}(t)\\ &= \Phi_2^*(t)\left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-it\right)}\right)^{k^*}\prod_{j=1}^{p}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\\ &= \Phi_2^*(t)\prod_{j=1}^{p-k^*}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\prod_{j=p-k^*+1}^{p}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n-1}{2}-it\right)}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\\ &= \Phi_2^*(t)\prod_{j=1}^{p-k^*}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\prod_{j=p-k^*+1}^{p}\frac{\Gamma\left(\frac{n-1}{2}+\frac{1}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n-1}{2}-it\right)}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-1}{2}+\frac{1}{2}+\frac{j-1}{p}-it\right)}\\ &= \prod_{j=p-k^*+1}^{p}\frac{\Gamma\left(\frac{n-1}{2}+1\right)\Gamma\left(\frac{n-1}{2}-it\right)}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-1}{2}+1-it\right)}\,\frac{\Gamma\left(\frac{n-1}{2}+\frac{1}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n-1}{2}+1-it\right)}{\Gamma\left(\frac{n-1}{2}+1\right)\Gamma\left(\frac{n-1}{2}+\frac{1}{2}+\frac{j-1}{p}-it\right)}\\ &\quad\times\prod_{j=1}^{p-k^*}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\times\Phi_2^*(t)\\ &= \left(\left(\frac{n-1}{2}\right)\left(\frac{n-1}{2}-it\right)^{-1}\right)^{k^*}\times\Phi_2^*(t)\\ &\quad\times\prod_{j=p-k^*+1}^{p}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n+1}{2}-it\right)}{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\prod_{j=1}^{p-k^*}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\\ &= \Phi_2^{**}(t)\,\Phi_1^{**}(t),\end{aligned}$$

with $\Phi_1^{**}(t)$ and $\Phi_2^{**}(t)$ in (17). $\square$

Proof of Theorem 5. The proof of this theorem is, in all respects, similar to the proof of Theorem 2. But since the depth of the GNIG distributions involved is different, we will still briefly state the proof.

If in the characteristic function of W in (17) we replace $\Phi_1^{**}(t)$ by $\sum_{k=1}^{h/2}\theta_k\,\mu^{\delta_k}(\mu-\mathrm{i}t)^{-\delta_k}$ we obtain

$$\Phi_W(t)\approx\sum_{k=1}^{h/2}\theta_k\,\mu^{\delta_k}(\mu-\mathrm{i}t)^{-\delta_k}\prod_{j=1}^{p-1}\left(\frac{n-j}{2}\right)^{\left\lfloor\frac{p-j+1}{2}\right\rfloor}\left(\frac{n-j}{2}-\mathrm{i}t\right)^{\left\lfloor\frac{p-j+1}{2}\right\rfloor},$$

that is the characteristic function of the mixture of $h/2$ GNIG distributions of depth p with cdf given in (18). The parameters δ_k, μ and θ_k are determined in such a way that (13) holds with $\Phi_1^{*}(t)$ replaced by $\Phi_1^{**}(t)$, which requires the evaluation of these parameters as the numerical solution of a system of h equations in all similar to the ones in (16), with $\Phi_1^{*}(t)$ replaced by $\Phi_1^{**}(t)$, leading to a near-exact distribution that matches the first h exact moments of W. $\square$

References

[1] C.A. Coelho, The Generalized Integer Gamma distribution — A basis for distributions in Multivariate Statistics, Journal of Multivariate Analysis 64 (1998) 86–102.
[2] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, 3rd ed., J. Wiley & Sons, New York, 2003.
[3] C.A. Coelho, The Generalized Near-Integer Gamma distribution: A basis for 'near-exact' approximations to the distributions of statistics which are the product of an odd number of independent Beta random variables, Journal of Multivariate Analysis 89 (2004) 191–218.
[4] R.P. Alberto, C.A. Coelho, Study of the quality of several asymptotic and near-exact approximations based on moments for the distribution of the Wilks Lambda statistic, Journal of Statistical Planning and Inference 137 (2007) 1612–1626.
[5] L.M. Grilo, C.A. Coelho, Development and study of two near-exact approximations to the distribution of the product of an odd number of independent Beta random variables, Journal of Statistical Planning and Inference 137 (2007) 1560–1575.
[6] F.J. Marques, C.A. Coelho, Near-exact distributions for the sphericity likelihood ratio test statistic, Journal of Statistical Planning and Inference 138 (2008) 726–741.
[7] C.A. Coelho, R.P. Alberto, L.M. Grilo, A mixture of Generalized Integer Gamma distributions as the exact distribution of the product of an odd number of independent Beta random variables, Journal of Interdisciplinary Mathematics 9 (2006) 229–248.
[8] C.A. Coelho, The wrapped Gamma distribution and wrapped sums and linear combinations of independent Gamma and Laplace distributions, Journal of Statistical Theory and Practice 1 (2007) 1–29.
[9] G.S. Mudholkar, M.C. Trivedi, C.T. Lin, An approximation to the distribution of the likelihood ratio statistic for testing the complete independence, Technometrics 24 (1982) 139–143.
[10] G.E.P. Box, A general distribution theory for a class of likelihood criteria, Biometrika 36 (1949) 317–346.
[11] M.S. Bartlett, A note on multiplying factors for various χ^2 approximations, Journal of the Royal Statistical Society, Ser. B 16 (1954) 296–298.
[12] J. Gil-Pelaez, Note on the inversion theorem, Biometrika 38 (1951) 481–482.